tual equivalence to 'hassled'. **CONCLUSIONS:** Though not always possible to translate an English word exactly into the target language, a full translation and linguistic validation process, including creation of a concept elaboration document followed by an in-depth discussion at back-translation review stage, enables a conceptually equivalent translation to be found. The concept elaboration document should, where possible, be created in conjunction with the instrument developer.

PRM41

IS IT TIME TO ELIMINATE THE ICER? USING NET BENEFITS TO REPORT THE RESULTS OF DETERMINISTIC COST-EFFECTIVENESS ANALYSES

O'Day K, McLaughlin T, Bramley T
Xcenda, LLC., Palm Harbor, FL, USA

BACKGROUND: Incremental cost-effectiveness ratios (ICERs) are used to report the results of cost-effectiveness (CE) analyses and represent the cost per unit of effectiveness of a more costly and more effective option. However, numerous conceptual and practical problems limit the usefulness of ICERs for decision making. These problems include, but are not limited to, negative ICERs, one-way sensitivity analyses, complexity of multiple comparator analyses, and statistical limitations with ratios. The net benefits approach was developed to address the statistical limitations of ICERs and is now an accepted methodology used in probabilistic sensitivity analysis to estimate CE confidence intervals and plot acceptability curves. However, despite the remaining challenges and limitations the use of ICERs persists, raising the question: Is it time to eliminate the use of ICERs in the reporting of CE analyses? **METHODS:** We propose expanding the net benefit method to present deterministic CE analysis results using a net monetary benefit (NMB) chart or table. A NMB chart is plotted with the x-axis representing the WTP threshold and the y-axis representing the NMB. The NMB of each option is a line with the intercept representing the cost and the slope representing the effectiveness across a specified range of WTP values. The line with the greatest NMB at a given WTP represents the most cost-effective option at that WTP. The vertical distance between two lines represents the incremental NMB. Dominance, extended dominance, and the frontier are captured graphically and intuitively. Multiple comparator analyses are simplified and one-way sensitivity analyses are enhanced due to the elimination of negative ICERs. **CONCLUSIONS:** A net benefits approach provides a more intuitive, informative, and useful method to present CE results than the use of ICERs. Moreover, it has the benefit of facilitating a uniform and consistent approach to presenting the results of deterministic and probabilistic CE analyses.

PRM42

VISUALLY EVALUATING THE MEASUREMENT COMPARABILITY BETWEEN PAPER-BASED AND ALTERNATE VERSIONS OF ADMINISTRATION OF THE LUNG FUNCTION QUESTIONNAIRE

Gilligan T[1], Nelson L[1], McLeod L[1], Dalal AA[2]
[1]*RTI Health Solutions, Research Triangle Park, NC, USA,* [2]*GlaxoSmithKline, Durham, NC, USA*

In randomized crossover designs, intraclass correlation coefficients (ICCs) are often used to assess the concordance between scores on different administration versions of patient-reported outcome (PRO) measures. An ICC and its associated criterion for "adequate" concordance enable analysts to simplify information and provide researchers with a quick and easy way to interpret analysis output. This strength of the ICC—its simplicity—may also be a weakness. Analysts may overlook important information (e.g., biases, outliers) when ICCs are used as the primary method for assessing concordance. One way to avoid overlooking important information is to include the evaluation of Bland-Altman plots when assessing concordance. Bland-Altman plots allow one to visually determine whether two measures produce similar scores, therefore, supplementing the concordance information gained from the ICC evaluation. ICCs and Bland-Altman plots complement each other's strengths. ICCs provide an efficient and concise estimate to determine the comparability of versions, while Bland-Altman plots provide a greater level of detail that incorporates a broader view of the analyzed distributions. The use of the two methods together provides a more holistic view of concordance. We present Bland-Altman plots and corresponding ICCs under a randomized crossover-design, using the Lung Function Questionnaire, a PRO instrument originally designed to be administered via paper, and later via three alternate administration versions (Web, interactive voice response system, and interview). We provide examples to illustrate instances in which ICCs and Bland-Altman plots agree and disagree. GSK study number: ADC001HO.

PRM43

BIOSIMILARS: DEMONSTRATING SIMILARITY THROUGH EVIDENCE

Kleintjens J[1], Dahal D[1], Mai JC[1], Doyle JJ[2]
[1]*Quintiles, Hawthorne, NY, USA,* [2]*Columbia University, New York, NY, USA*

OBJECTIVES: The differences in the active substance of biosimilars compared to their originator reference product can cause risks that are unique to biologics, mainly: immunogenicity, long-term safety risks, and lack of efficacy. These risks are unknown at the launch of a biosimilar and can lead to unexpected costs for payers. The aim of this abstract is to describe a methodology for evaluating the unknown risks of biosimilars. **METHODS:** A structured literature review revealed that for many biologics, post-marketing observational studies have been set up to identify long-term safety and efficacy outcomes. These studies are a useful source of information to quantify the unknown risks of biosimilars and define methods of minimizing those risks. The information required is product- and population-specific. This information first includes potential safety issues such as immunogenicity (all biologics), serious infections and autoimmune disorders (anti-TNFs, interferons), and increased mortality and cardiovascular events (epoetin). Second, information is available on long-term benefits such as clinical outcomes that improve overall survival (e.g. reduced recurrence of malignancies through interferon use and reduced cardiovascular events through insulin use), reduction in healthcare resource utilization (epoetin, somatropin), and proportion of long-term responders (figrastim, anti-TNF, somatropin). Finally, observational data can be used to optimize treatment regimens to achieve maximum treatment benefit (epoetin, insulin, somatropin). All these data can be used in an economic evaluation where the unknown risks for biosimilars are quantified through worst- and best-case scenarios. **CONCLUSIONS:** Often, payers are attracted to biosimilars that have the lowest acquisition costs. However, the risks of unknown information for these biosimilars should be valued against the lower price of these drugs. Observational data for the originator biologic product can be leveraged to quantify these risks. This can help determine for which drugs and for which populations the unknown risks outweigh the reduction in acquisition costs.

PRM44

OPTIMUM METHODS FOR PRO TRANSLATION AND LINGUISTIC VALIDATION METHODOLOGY: ONE BACK TRANSLATION OR TWO?

Furtado T, Anderson H, Griffin A, Wild D
Oxford Outcomes Ltd., Oxford, UK

OBJECTIVES: The translation of PRO measures requires a rigorous procedure, including dual forward translations, reconciliation, back translation and review, and debriefing interviews. The relevance of including blinded back translations has been widely discussed; however, there has been little discussion around back translation methods. This research aims to gauge the importance of having two back translations versus only one. **METHODS:** Past translation and linguistic validation projects employing the procedures outlined above were reviewed to compare the methodology of using one back translation versus two. **RESULTS:** In the dual-back projects, numerous instances were found in which only one back translator detected an issue. For example: 1. Simple mistranslations can be revealed by one translator but not another; e.g. 'activities at home' was back translated verbatim by one back translator but as 'household activities' by another, revealing that the translated term was too narrow and related only to chores; 2. Similarly, contextual mistranslations may become apparent in dual back translations. In one ePRO script, 'enter training module' meant to click through to the next page. One back translator wrote 'enter', while 'insert' in the second translation highlighted that the wrong term had been used in this context; 3. Dual translations are also useful for elucidating nuances in meaning. In Danish, the phrase 'bad tempered' became 'lose my temper' in one back translation, allowing the lead translator to alter the ambiguous translation; 4. Dual meanings in the target language may also be highlighted, for example in Gujarati 'hospitalisations' was correctly back translated by one translator, but the other translated it as 'clinic', showing an ambiguity in the translation. **CONCLUSIONS:** The high proportion of issues highlighted by only one back translator, show the importance of using two blinded back translators in the translation of PRO measures.

PRM45

PILOT TESTING TRANSLATIONS OF PRO MEASURES WITH SENSITIVE POPULATIONS

Verjee-Lorenz A, Clayson D, Two R
PharmaQuest Ltd, Banbury, Oxfordshire, UK

The ISPOR Principles of Good Practice paper on the translation of patient-reported outcome (PRO) measures includes cognitive debriefing as a key step in the translation process. Cognitive debriefing refers to the process of asking patients to describe to an interviewer what each question/instruction means to them in their own words. This allows experienced project managers to then determine whether the patient has correctly understood the conceptual meaning of the question/instruction and therefore, by implication, whether the translation is accurate. Despite the FDA guidelines emphasising the importance of assessing the content validity of PRO translations, cognitive debriefing is not always feasible or ethically acceptable with certain patient groups. These groups include children, severely ill patients and patients with mental health problems, for whom prolonged interviews could cause distress. In these circumstances, alternative methodologies should be employed to establish whether the wording is suitable for the given population. Evidence from working with such patients indicates that clinician reviews, caregiver reviews and assessments of language complexity and suitability are all useful alternative methods for establishing the acceptability of PRO translations for particular patient populations. Two examples include pilot testing mental health measures with clinicians as a substitute for patients experiencing acute symptoms and pilot testing paediatric measures with health practitioners and teachers working with children from the target age group. Cognitive debriefing is a useful tool in the translation and linguistic validation of PRO measures but a more flexible approach is required to ensure that certain patients are not unduly distressed and burdened by the process.

PRM46

R . . . YOU AWARE HOW USEFUL IT IS? THE VALUE OF CORRELATION COEFFICIENTS IN META-ANALYSIS

Al-Dakkak I, Patel S, Jen MH, von Maltzahn R
HERON Evidence Development Ltd, Luton, UK

Meta-analysis involves pooling effect sizes to combine results from studies attempting to answer similar research questions. Typically, a common metric is used to estimate associations between independent and dependent variables in pooled studies. Nonetheless, studies vary considerably in their measurement of effect sizes as well as the nature of studied variables. The calculation of an effect size *r* allows the pooling of results that are reported in a variety of forms. In the presence